

## (12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

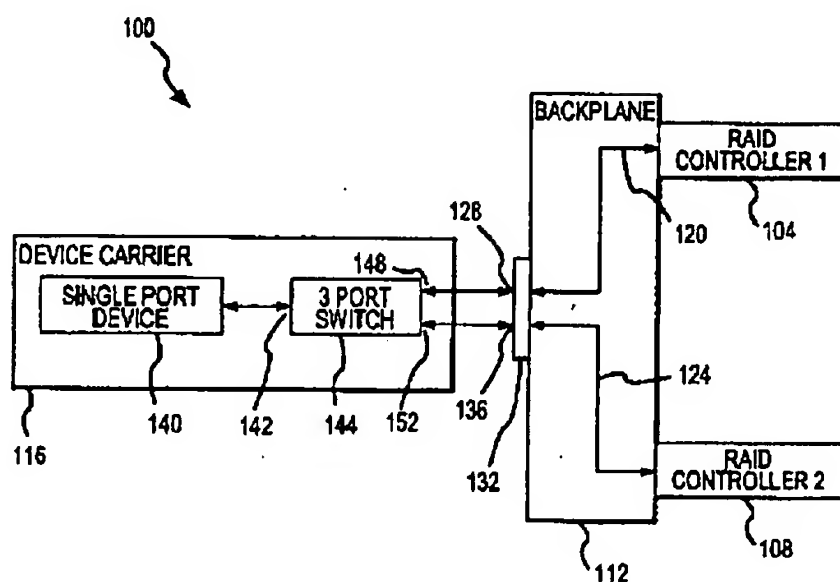
(19) World Intellectual Property Organization  
International Bureau(43) International Publication Date  
6 November 2003 (06.11.2003)

PCT

(10) International Publication Number  
WO 03/091887 A1

- (51) International Patent Classification<sup>7</sup>: G06F 13/00
- (21) International Application Number: PCT/US03/11905
- (22) International Filing Date: 16 April 2003 (16.04.2003)
- (25) Filing Language: English
- (26) Publication Language: English
- (30) Priority Data:  
10/131,387 23 April 2002 (23.04.2002) US
- (71) Applicant: ADAPTEC, INC. [US/US]; 691 South Milpitas Blvd., Milpitas, CA 95035 (US).
- (81) Designated States (*national*): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EB, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NI, NO, NZ, OM, PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, TJ, TM, TN, TR, TT, TZ, UA, UG, UZ, VC, VN, YU, ZA, ZM, ZW.
- (84) Designated States (*regional*): ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IT, LU, MC, NL, PT, RO, SE, SI, SK, TR), OAPI patent (BF, BJ, CF, CG, CL, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).
- (72) Inventor: LYNN, William, E.; 846 Atwood Street, Longmont, CO 80501 (US).
- (74) Agents: KNEPPER, Bradley, M. et al.; Sheridan Ross P.C., 1560 Broadway, Suite 1200, Denver, CO 80202-5141 (US).
- Published:  
— with international search report
- For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

(54) Title: METHOD AND APPARATUS FOR DUAL PORTING A SINGLE PORT SERIAL ATA DISK DRIVE



(57) Abstract: A method and an apparatus for providing multiple ports (148, 152) to a single port device (140) are provided. According to the present invention, single port devices that are incapable of communicating across a shared communication bus, and therefore require a point to point communication channel, may be interconnected to multiple hosts using a switch assembly (144). The switch assembly selectively establishes a communication path between a single host (104 or 108) and a device (140). Additionally, the present invention provides a method and apparatus by which computer subsystems, such as RAID storage subsystems, may be implemented using inexpensive storage devices. Furthermore, the present invention provides a method and apparatus for providing a high reliability computer system by configuring the subsystem such that no single active component failure will result in the failure of the entire subsystem.

WO 03/091887

PCT/US03/11905

## METHOD AND APPARATUS FOR DUAL PORTING A SINGLE PORT SERIAL ATA DISK DRIVE

### FIELD OF THE INVENTION

The present invention relates to providing redundant computer systems. In particular,  
5 the present invention relates to the use of multiple hosts or controllers in connection with a single port device.

### BACKGROUND OF THE INVENTION

Computer systems require reliable storage for large amounts of data. Often,  
10 redundant arrays of independent (or inexpensive) disks (RAID) devices are used to provide such storage. In general, RAID devices involve storing data on a plurality of hard disk drives. The use of RAID techniques increases the reliability and/or speed of data storage and retrieval. Although there are various schemes, or RAID levels, according to which data may be stored and retrieved, RAID systems all generally involve the distribution of data across  
15 multiple storage devices. The distribution and retrieval of data from the devices is usually performed by a controller interconnected to each of the storage devices in the RAID system.

In order to further enhance the fault tolerant aspects of data storage, it is desirable to ensure that there is no single point of failure within a storage subsystem. Therefore, in addition to distributing data among a number of storage devices, high reliability storage  
20 systems have used two controllers, to allow the storage system to continue to operate even if a fault develops with respect to the primary controller. In such systems, the storage devices must be capable of passing data to and receiving data and commands from either of the controllers. Accordingly, high reliability storage systems have typically used small computer system interface (SCSI) or Fibre Channel type storage devices, that can be operated on a  
25 shared bus and/or provide multiple ports for interconnecting to controllers.

Although SCSI and fiber channel compliant devices facilitate the use of redundant controllers in a storage system, such devices are typically much more expensive than devices that are not capable of operating in connection with a shared communication bus. However, less expensive devices, such as those adhering to the advanced technology attachment (ATA)  
30 or serial ATA (SATA) standards must be interconnected with a controller in a point to point relationship (*i.e.* over a dedicated signal line). Accordingly, high reliability systems, such as RAID storage systems, using redundant device controllers have been required to use the more expensive SCSI or Fibre Channel devices.

WO 03/091887

PCT/US03/11905

2

Therefore, it would be desirable to provide for redundant or multiple hosts or controllers in connection with inexpensive devices. Furthermore, it would be advantageous to allow devices incapable of operation on a shared bus to be interconnected to a number of hosts or controllers in a way that is reliable, and that is itself relatively inexpensive to  
5 implement.

### SUMMARY OF THE INVENTION

In accordance with the present invention, a method and an apparatus for interconnecting single port devices to a plurality of hosts or controllers are provided. In  
10 general, the present invention provides a switch assembly that allows a single device to be selectively interconnected to one of a plurality of hosts.

According to an embodiment of the present invention, a first port of a three-port switch is interconnected to a single port device. The second port is interconnected to a first host, and the third port is interconnected to a second host. Accordingly, by operation of the  
15 switch, the single port device may be selectively interconnected to either the first or second host.

According to a further embodiment of the present invention, a switch assembly capable of operating in a single threaded mode is provided. According to such an embodiment, a switch assembly is operated to interconnect a first port associated with a  
20 device to a second port associated with a host in response to a command or data received at the second port of the switch assembly. The interconnection between the device and the first host may be maintained until a command or data is passed from the device to the first host in response to the command or data received at the second port. Similarly, the switch may interconnect the first port to a third port in response to a command or data received at the  
25 third port from a second host. The interconnection between the first and third ports may be maintained until a command or data is passed from the device to the second host through the switch assembly. According to a further embodiment of the present invention, if a command or data is received at either of the second or third ports while a command or data from the device in response to a command or data received at the other of the ports is pending or in  
30 progress, the host initiating the later arriving command or data will be signaled that the switch assembly is not ready to receive the command or data. Such a signal may be in the form of a failure to issue a ready signal to the host.

WO 03/091887

PCT/US03/11905

3

In accordance with another embodiment of the present invention, a switch assembly capable of operating in a queued environment is provided. According to such an embodiment, commands or data passed between a first host and a device may be associated with a first queue tag. Similarly, commands and data passed between a second host and the  
5 device may be associated with a second queue tag. A switch assembly for selectively interconnecting the device to either the first host or the second host operates to pass commands or data from the device associated with the first queue tag to the first device, while passing commands and data associated with the second queue tag to the second device.

Based on the foregoing summary, a number of salient features of the present  
10 invention are readily discerned. A method and an apparatus for interconnecting a single port device to a number of hosts are provided. Accordingly, the present invention allows for the use of redundant hosts in connection with relatively inexpensive, single port devices.

Additional advantages of the present invention will become readily apparent from the following discussion, particularly when taken together with the accompanying drawings.

15

#### BRIEF DESCRIPTION OF THE DRAWINGS

Fig. 1 is a block diagram depicting a single port device interconnected to a plurality of hosts in accordance with an embodiment of the present invention;

Fig. 2 is a block diagram depicting a plurality of single port devices interconnected  
20 to a plurality of hosts in accordance with an embodiment of the present invention;

Fig. 3 is a functional block diagram depicting a switch assembly in accordance with an embodiment of the present invention;

Fig. 4 is a flow chart depicting the operation of a subsystem in accordance with an embodiment of the present invention; and

Fig. 5 is a flow chart depicting the operation of a subsystem in accordance with  
25 another embodiment of the present invention.

#### DETAILED DESCRIPTION

With reference now to Fig. 1, a computer subsystem 100 in accordance with an  
30 embodiment of the present invention is depicted. The subsystem 100 generally includes first 104 and second 108 hosts, also referred to herein as initiators or controllers. In addition, the subsystem 100 includes a signal back plane 112, and a device carrier 116.

WO 03/091887

PCT/US03/11905

4

The signal back plane 112 provides a first back plane signal path 120, and a second back plane signal path 124. With respect to a first end of the first back plane signal path 120, the signal back plane 112 provides an interconnection to the first host 104. At a second end of the first back plane signal path 120, the signal back plane 112 provides a port 128, as part of a dual port connector 132. With respect to a first end of the second back plane signal path 124, the signal back plane 112 provides an interconnection to the second host 108. At a second end of the second back plane signal path 124, a port 136 is provided as part of the dual port connector 132.

The device carrier 116 generally includes a single port device 140 interconnected to a first port 142 of a switch assembly 144. The device carrier 116 may be configured to facilitate the interconnection of the device 140 and associated switch assembly 144 to the signal back plane 112. For example, the device carrier 116 may be configured to be slidably received by an enclosure or rack interconnected to or incorporating the signal backplane 112 and provided as part of a subsystem 100, such as a RAID storage system. Furthermore, by positioning the device carrier 116 in a receiving slot of an enclosure or rack, an interconnection between second 148 and third 152 ports, and corresponding ports 128, 136 on the dual port connector 132 of the signal back plane 112, may be established. The subsystem 100 may also be implemented without an enclosure or rack.

The single port device 140 may include any device that requires a point to point, dedicated signal path for communications. That is, the single port device 140 can be a device that is incapable of operation in connection with a shared communication bus, and that provides a single port for interconnecting the device 140 to a communication channel.

The switch assembly 144 selectively establishes a signal path between the first port 142 and either the second port 148 or the third port 152 of the switch assembly 144. In the subsystem 100 illustrated in Fig. 1, it can therefore be appreciated that when the switch assembly 144 is in a first operating mode, the first port 142 is interconnected to the second port 148 of the switch assembly 144. So configured, a communication channel or signal path is established between the single port device 140 and the first host 104. When the switch assembly 144 is in a second operating mode, the first port 142 is interconnected to the third port 152. In this second operating mode, the switch assembly 144 enables a signal or communication path between the single port device 140 and the second host 108.

WO 03/091887

PCT/US03/11905

5

Accordingly, it can be appreciated that the subsystem 100 allows two different hosts 104, 108 to be interconnected to a single port device 140. Furthermore, it can be appreciated that the subsystem 100 allows all active components (e.g., the hosts 104, 108, devices 140, and switch assemblies 144) associated with the system 100 to be selectively interconnected to the subsystem 100, and is configured such that no one active component could cause the entire subsystem 100 to fail.

With reference now to Fig. 2, a computer subsystem 200 in accordance with another embodiment of the present invention is illustrated. The subsystem 200 interconnects a plurality of device carriers 216a, 216b, and 216n, each having an associated single port storage device 240, to first 204 and second 208 controllers. The subsystem 200 in Fig. 2 is therefore illustrative of a RAID storage subsystem with redundant RAID controllers 204, 208 and a plurality of storage devices 240. Accordingly, the signal back plane 212 of the system 200 provides a plurality of first backplane signal paths 220a, 220b, and 220n, and a plurality of second backplane signal paths 224a, 224b and 224n.

The storage devices 240 may comprise hard disk drives that require a point to point connection to an associated controller. For example, the storage devices 240 may comprise advanced technology attachment (ATA) or serial ATA (SATA) devices. Each device 240 may be selectively interconnected to either the first 204 or the second controller 208. In particular, the switch assemblies 244 operate to selectively interconnect an associated device 240 to the first controller 204. For example, when a first switch assembly 244a is in a first mode of operation, a first device 240a may be interconnected to the first controller 204 via a signal path comprising the first backplane signal path 220a, the second port 248a of the switch assembly 244a, and the first port 242a of the switch assembly 244a. When the switch assembly 244a is in a second mode of operation, the device 240a may be interconnected to the second controller 208 via a signal path comprising the second backplane signal path 224a, the third port 252a of the switch assembly 244a, and the first port 242a of the switch assembly 244a.

As depicted in Fig. 2, the signal backplane 212 and the device carriers 216 may be interconnected to a rack 256 when the system 200 is operative. Also illustrated in Fig. 2 is a communication bus or computer network 260 interconnecting the subsystem 200 to a plurality of computers 264. Accordingly, Fig. 2 also illustrates an example computer system

WO 03/091887

PCT/US03/11905

6

264 comprising a computer storage subsystem 200 in accordance with an embodiment of the present invention. As can be appreciated by one of skill in the art, the communication bus or computer network 260 may include any network suitable for interconnecting computers and computing devices, such as an ethernet network. Furthermore, the computer network  
5 260 may comprise a communications bus used in connection with communications involving peripheral computer devices, such as a peripheral component interconnect (PCI) communication bus.

The computer subsystem illustrated in Fig. 2 provides a RAID storage subsystem that protects against data loss due to a single point of failure. In particular, the subsystem 200  
10 illustrated in Fig. 2 does not include any active devices that would render the subsystem 200 inoperative if any one device were to fail. For example, the inclusion of first 204 and second 208 RAID controllers provides redundancy should the primary controller (e.g., the first controller 204) fail. In addition, the data is distributed among the devices 240 such that the loss of any one device does not result in a loss of the data stored on that device 240. To  
15 facilitate the replacement of a failed device 240 or switch assembly 244, each device 240 and its associated switch assembly 244 is interconnected to an associated device carrier 216 that can be removed from the rack 256 and replaced without resulting in a total failure of the subsystem 200. Furthermore, it will be noted that the backplane 212 contains no active components, removing the backplane 212 as a potential source of system failure.

20 With reference now to Fig. 3, a switch assembly 144, 244 is depicted in functional block diagram form. As shown in Fig. 3, the switch assembly 144, 244 comprises a first port 142, 242, a second port 148, 248, and a third port 152, 252. The first port 142, 242 may be selectively interconnected to either the second port 148, 248 or the third port 152, 252 by a switch block 304. Switch logic 308 controls the operation of the switch block 304.

25 The various ports 142, 242, 148, 248, 152, 252 may comprise mechanical interconnections for mating the switch assembly 144, 244 to complementary connectors that are, for example, included as part of the backplane 112, 212, and/or as part of the device carrier 116, 216. The switch 304 may provide a logical or physical interconnection between the first port 142, 242 and either the second port 148, 248 or the third port 152, 252. The  
30 switch logic 308 may comprise a simple state machine implemented using a controller or a simple programmable logic device.

WO 03/091887

PCT/US03/11905

7

With reference now to Fig. 4, a flow chart depicting the operation of a system 100, 200 in accordance with an embodiment of the present invention is illustrated. In particular, the process steps illustrated in Fig. 4 may be considered with respect to a single switch assembly 144, 244 (e.g., switch assembly 244a in Fig. 2). From an initial starting point, labeled A in Fig. 4, a determination is made as to whether a command or data was received at the second port 148, 248 of the switch assembly 144, 244 (step 404). If a command or data was received at the second port 148, 248, a determination is made as to whether a transfer of a command or data between the first port 142, 242 and the third port 152, 252 of the switch assembly 144, 244 is pending or in progress (step 408). If such a transfer is pending or in progress, the host interconnected to the second port 148, 248 (i.e. the first host 104, 204) is notified that the system 100, 200 is not ready to pass the command or data received at the second port 148, 248 to the device 140, 240 (step 412). Such notification may be in the form of a failure to issue a ready signal from the switch assembly 144, 244 to the first controller 204. The system may then return to point A.

If at step 408 it is determined that no transfer of a command or data between the first port 142, 242 and the third port 152, 252 is pending or in progress, a signal path is established between the first 142, 242 and second 148, 248 ports (step 416). That is, the switch assembly 144, 244 is operated to interconnect the first port 142, 242 to the second port 148, 248. The command or data may then be passed from the second port 148, 248 to the first port 142, 242 (step 420). The system may then return to point A.

If a command or data is not received at the second port 148, 248, a determination is made as to whether a command or data has been received at the third port 152, 252 (step 424). If a command or data was received at the third port 152, 252, a determination is made as to whether a transfer of a command or data between the first port 142, 242 and second port 148, 248 is pending or in progress (step 428). If such a transfer is pending, the host connected to the third port 152, 252 (i.e. the second host 108, 208) is notified that the third port 152, 252 is not ready (step 432). Such notification may be in the form of a failure to issue a ready signal to the host 108, 208. The system may then return to point A.

If a transfer of a command or data between the first port 142, 242 and the second port 148, 248 is not pending or in progress, a signal path is established between the first 142, 242



WO 03/091887

PCT/US03/11905

8

and third 152, 252 ports (step 436). The command or data is then passed from the third port 152, 252 to the first port 142, 242 (step 440). The system may then return to point A.

If at step 424 it is determined that a command or data has not been received at the third port 152, 252, a determination is made as to whether a command or data was received at the first port 142, 242 (step 444) (i.e. from a device 140, 240). If no such command or data was received, the system returns to point A. If a command or data was received at the first port 152, 252, the command or data is sent through the switch assembly 144, 244 over the signal path last established by the switch assembly 144, 244 between the first 142, 242 and either of the second 148, 248 or third 152, 252 ports (step 448). That is, if the previous transfer of a command or data handled by the switch assembly was made between the second port 148, 248 and first port 142, 242, the command or data received at the first port 142, 242 will be routed to the second port 148, 248. Conversely, if the previous transfer of a command or data handled by the switch assembly 144, 244 was made between the third port 152, 252 and the first port 142, 242, then the command or data received at the first port 142, 242 is passed to the third port 152, 252. At step 452, the connection between the first port 142, 242 and the second 148, 248 or third ports 152, 252 is torn down. The system may then return to point A.

From the description given above, it can be appreciated that an embodiment of the present invention maintains an interconnection between a device 140, 240 and a controller 104, 204 or 108, 208 until a command or data is received in reply at the first port 142, 242 in reply to the previous command or data received at the second port 148, 248 or third port 152, 252. Accordingly, a command or data received at the first port 142, 242 of a switch assembly 144, 244 is passed to the port that received the previous command or data from a controller 104, 204 or 108, 208.

With reference now to Fig. 5, a flow chart depicting the operation of a subsystem 100, 200 in accordance with another embodiment of the present invention is illustrated. Initially, at step 500, a command or data is received at a port of a switch assembly 144, 244. At step 504, a determination is made as to whether the command or data was received at the second port 148, 248 or the third port 152, 252 of the switch assembly 144, 244. If the command or data was received at the second port 148, 248 or the third port 152, 252, the command or data

WO 03/091887

PCT/US03/11905

9

is passed to the first port 142, 242 of the switch assembly (step 508). The system then returns to step 500.

If the command or data was not received at the second port 148, 248 or the third port 152, 252, a determination is made as to whether the command or data was received at the first port 142, 242 (*i.e.* from a device 140, 240) (step 510). If the command or data was received at the first port 142, 242, a determination is made as to whether a queue tag received with the command or data associates that command or data with the second port 148, 248 (step 512). If the queue tag does associate the command or data with the second port 148, 248, that command or data is passed to the second port 148, 248 (step 516). The system then returns to step 500.

If a queue tag received with the command or data at the first port 142, 242 associates that command or data with the third port 152, 252 (step 520) the command or data is passed to the third port 152, 252 (step 524). The system then returns to step 500. If no queue tag associating the command or data with the third port 152, 252 is detected at step 520, the system returns to step 500.

From the description set forth above with respect to Fig. 5, it can be appreciated that an embodiment of the present invention may utilize queue tags to direct commands or data received at the first port 142, 242 to either the second port 148, 248 or the third port 152, 252. Furthermore, it can be appreciated that the device 140, 240 is not required to recognize queue tags. Instead, a queue tag provided together with a command or data is recognized by the switch logic 308 and an appropriate path through the switch assembly 144, 244 is established.

If a command or data is received at any one of the ports provided as part of a switch assembly 144, 244 while a data or command is being transmitted between two other of the ports of the switch assembly 144, 244, a not ready signal may be provided by the switch assembly 144, 244 to the host 104, 204, 108, 208 or device 140, 240 that sent the command or data to the switch assembly 144, 244. In addition, it should be appreciated that with any embodiment of the present invention, provisions may be made for handling a command or data received at one of the ports of a switch assembly 144, 244 at the same time that a command or data is received at another one of the ports of a switch assembly 144, 244. For example, an order of preference may be established between the various ports of a switch

WO 03/091887

PCT/US03/11905

10

assembly 144, 244. In addition, it should be appreciated that the not ready signal described herein may comprise the absence of a ready signal. That is, the switch assembly 144, 244 may normally signal that it is ready to receive the command or data by providing a ready signal to the source of the command or data (e.g., a host 104, 204, 108, 208 or device 140, 240). However, if the switch assembly 144, 244 is not ready to receive a command or data, a ready signal may not be provided to the source of the command or data.

Although the foregoing describes use of the present invention in connection with ATA and SATA devices, and gives particular examples related to RAID subassemblies, it should be appreciated that the present invention is not so limited. For example, the present invention may be utilized in connection with providing multiple communication paths to multiple hosts for any device configured for communicating over a point to point communication channel. Furthermore, the present invention may be utilized in connection with providing multiple ports to multiple devices configured for interconnection to a host using a star topology. It will additionally be appreciated that the present invention is particularly useful in connection with the implementation of RAID subsystems utilizing inexpensive (e.g., ATA or SATA) hard disk drives. Furthermore, it should be appreciated that a switch assembly 144, 244 is not limited to having three communication ports. Rather, a switch assembly in accordance the present invention may have  $n$  ports, allowing  $n-1$  hosts to be interconnected to a device, where  $n$  is equal to or greater than 3.

The foregoing discussion of the invention has been presented for purposes of illustration and description. Further, the description is not intended to limit the invention to the form disclosed herein. Consequently, variations and modifications commensurate with the above teachings, within the skill and knowledge of the relevant art, are within the scope of the present invention. The embodiments described hereinabove are further intended to explain the best mode presently known of practicing the invention and to enable others skilled in the art to utilize the invention in such or in other embodiments and with various modifications required by their particular application or use of the invention. It is intended that the appended claims be construed to include the alternative embodiments to the extent permitted by the prior art.

WO 03/091887

PCT/US03/11905

11

What is claimed is:

1. An apparatus for providing a plurality of ports to a device using a point to point communications protocol, comprising:
  - a first data port, wherein said first data port is adapted for interconnection to said device;
  - a second data port;
  - a first signal path, wherein said second data port may be selectively interconnected to said first data port by said first signal path;
  - a third data port; and
  - a second signal path, wherein said third data port may be selectively interconnected to said first data port by said second signal path, wherein data is passed over said first signal path in a first mode of operation, and wherein data is passed over said second signal path in a second mode of operation.
2. The apparatus of Claim 1, wherein said first mode of operation is entered in response to at least one of a command and data received at said second data port, wherein at least one of a command and data is prevented from being passed over said second signal path until at least one of a command and data is received from said device at said first data port.
3. The apparatus of Claim 1, wherein a queue tag is associated with commands and data passed over said first and second signal paths, wherein communications having a first queue tag value are passed over said first signal path, and wherein communications having a second queue tag value are passed over said second signal path.
4. The apparatus of Claim 1, further comprising:
  - a first host interconnected to said first signal path; and
  - a second host interconnected to said second signal path, wherein at least one of a command and data is passed between said device and said first host over said first signal path in said first mode of operation, and wherein at least one of a command and data is passed between said device and said second host over said second signal path in said second mode of operation.
5. The apparatus of Claim 4, wherein said first host comprises a first controller and wherein said second host comprises a second controller.
6. The apparatus of Claim 4, wherein said first host comprises a first RAID controller, and wherein said second host comprises a second RAID controller.

PCT/US03/11905

WO 03/091887

12

7. The apparatus of Claim 1, wherein said device comprises a storage device.
8. The apparatus of Claim 1, wherein said device comprises a single port device.
9. The apparatus of Claim 1, wherein said device comprises an SATA device.
10. The apparatus of Claim 4, further comprising a RAID system back plane
- 5 wherein said first signal path is interconnected to said first host by a first back plane signal path, and wherein said second signal path is interconnected to said second host by a second backplane signal path.
11. A data storage system, comprising:
- a first device controller;
- 10 a second device controller;
- a plurality of data storage devices;
- a plurality of switch assemblies, wherein each of said switches comprises at least first, second, and third ports, wherein said switch assemblies operate to selectively interconnect said first port to one of said at least second and third ports, and wherein a first of said ports
- 15 is interconnected to a one of said data storage devices;
- a backplane, comprising:
- interconnections to said at least second and said third ports of said switch assemblies; and
- an interconnection to said first device controller;
- 20 an interconnection to said second device controller;
- a plurality of first backplane signal paths, wherein said first device controller may communicate with a second port of any of said switch assemblies; and
- a plurality of second backplane signal paths, wherein said second device controller may communicate with a third port of any of said switch assemblies.
- 25 12. The data storage system of Claim 11, wherein said first device controller comprises a RAID controller, and wherein said second device controller comprises a RAID controller.
13. The data storage system of Claim 11, wherein said data storage devices comprise SATA data storage devices.
- 30 14. The data storage system of Claim 11, further comprising a plurality of carriers, wherein a one of said data storage devices and a one of said switches is interconnected to each of said carriers.

WO 03/091887

PCT/US03/11905

13

15. A method for providing point to point communications between a plurality of hosts and a device, comprising:

establishing a first communication channel between said device and a first of said hosts in response to a communication received from said first host;

5 passing at least one of a command and data between said device and said first host over said first communication channel; and

establishing a second communication channel between said device and a second of said hosts in response to a communication received from said second host, wherein said device is incapable of operation on a shared bus.

10 16. The method of Claim 15, further comprising:

preventing establishment of a communication channel between said device and said second host in response to a communication received from said second host while a transfer of at least one of a command and data over said communication channel between said device and said first host is pending.

15 17. The method of Claim 15, further comprising:

passing at least one of a command and data from said device to said first host over said first communications channel.

18. The method of Claim 15, wherein said first and second communications channels are established by a switch assembly having a plurality of states, comprising:

20 a first state, wherein communications between said device and said first host are enabled, and wherein communications between said device and said second host are disabled; and

a second state, wherein communication between said device and said second host are enabled, and wherein communications between said device and said first host are disabled.

25 19. The method of Claim 18, wherein said states of said switch assembly further comprise:

a third state, wherein no communications between said device and said first and second hosts are enabled.

30 20. The method of Claim 18, wherein said first state is maintained until no transfer of at least one of a command and data between said device and said first host are pending.

WO 03/091887

PCT/US03/11905

14

21. The method of Claim 18, wherein said switch assembly enters said second state in response to a communication from said second host.

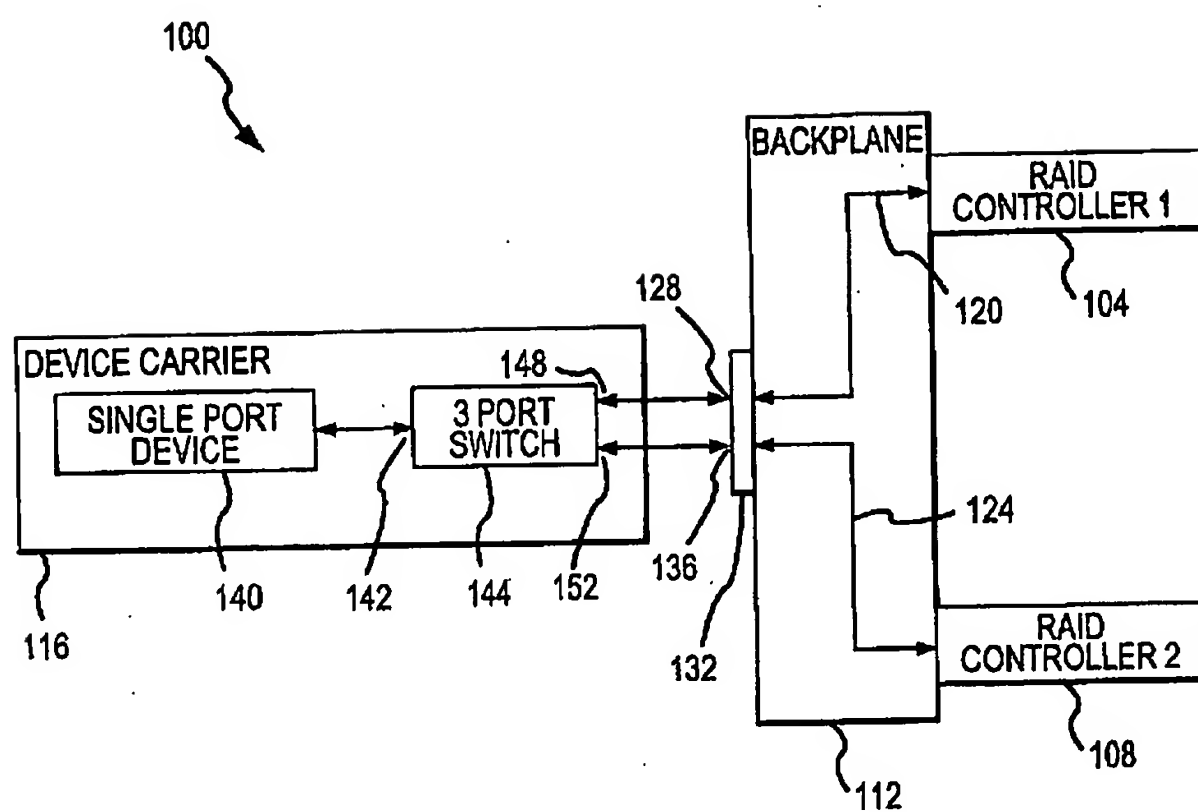
22. The method of Claim 18, wherein in response to at least one of a command and data received at said switch assembly from said second host while said switch assembly is in said first state and while a transfer of at least one of a command and data between said device and said first host is pending no ready signal is passed from said switch assembly to said second host, wherein said at least one of a command and data from said second host is not passed to said device.

23. The method of Claim 15, wherein at least one of a command and data having an associated first queue tag is passed between said device and said first host, and wherein at least one of a command and data having an associated second queue tag is passed between said device and said second host.

WO 03/091887

PCT/US03/11905

1/5





WO 03/091887

PCT/US03/11905

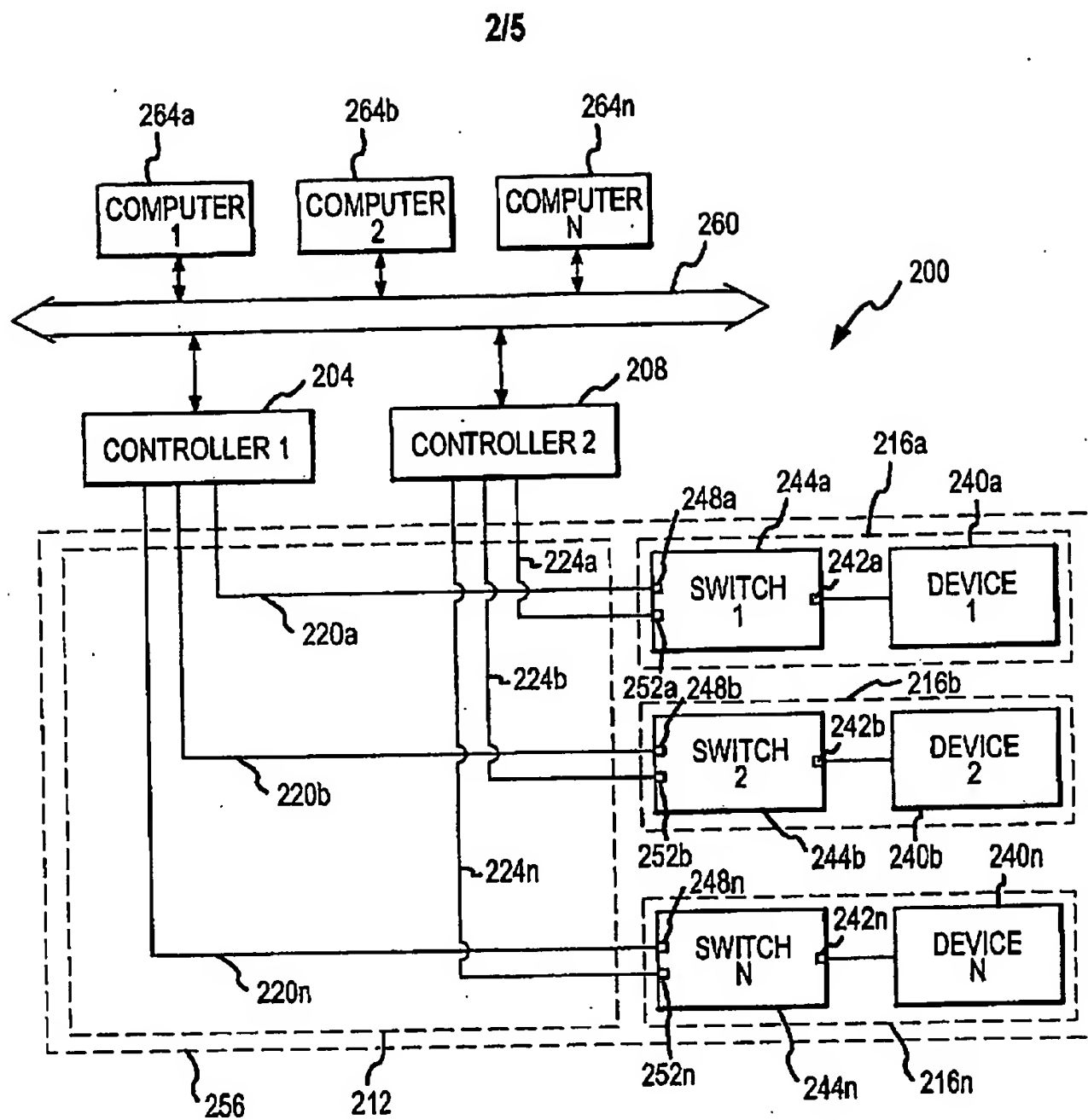


FIG.2

WO 03/091887

PCT/US03/11905

3/5

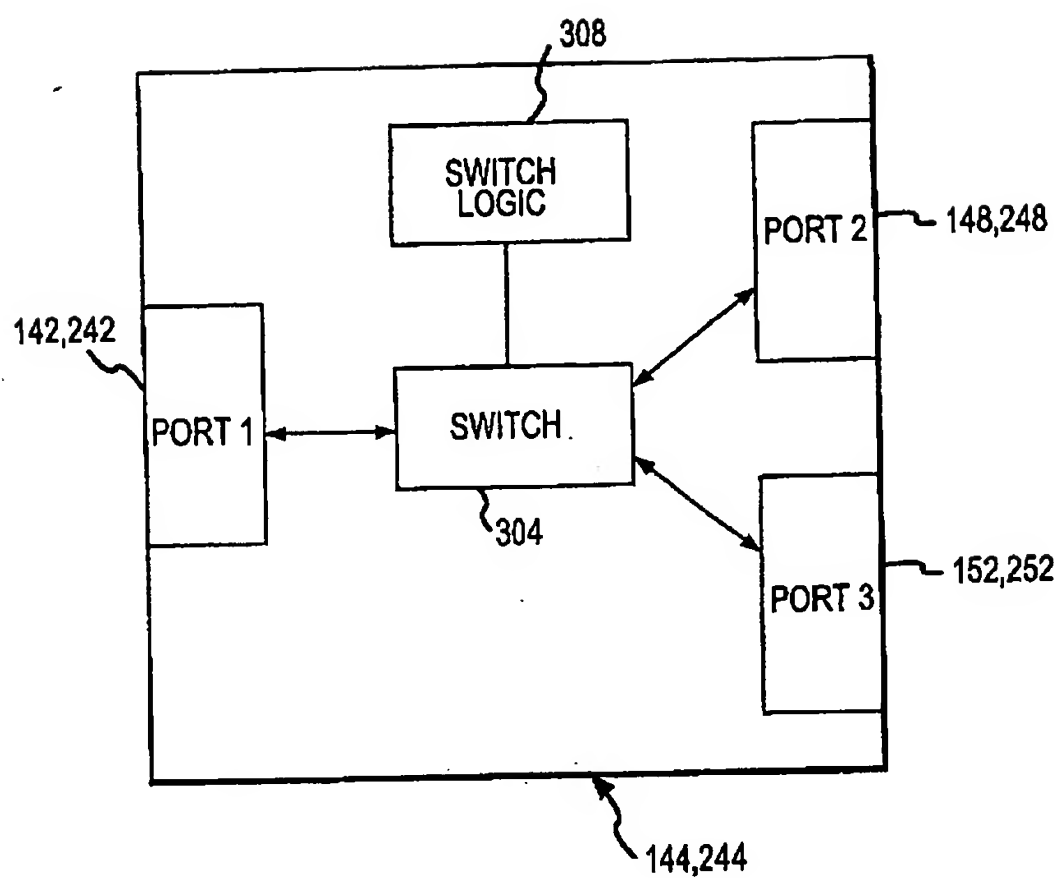


FIG.3

WO 03/091887

PCT/US03/11905

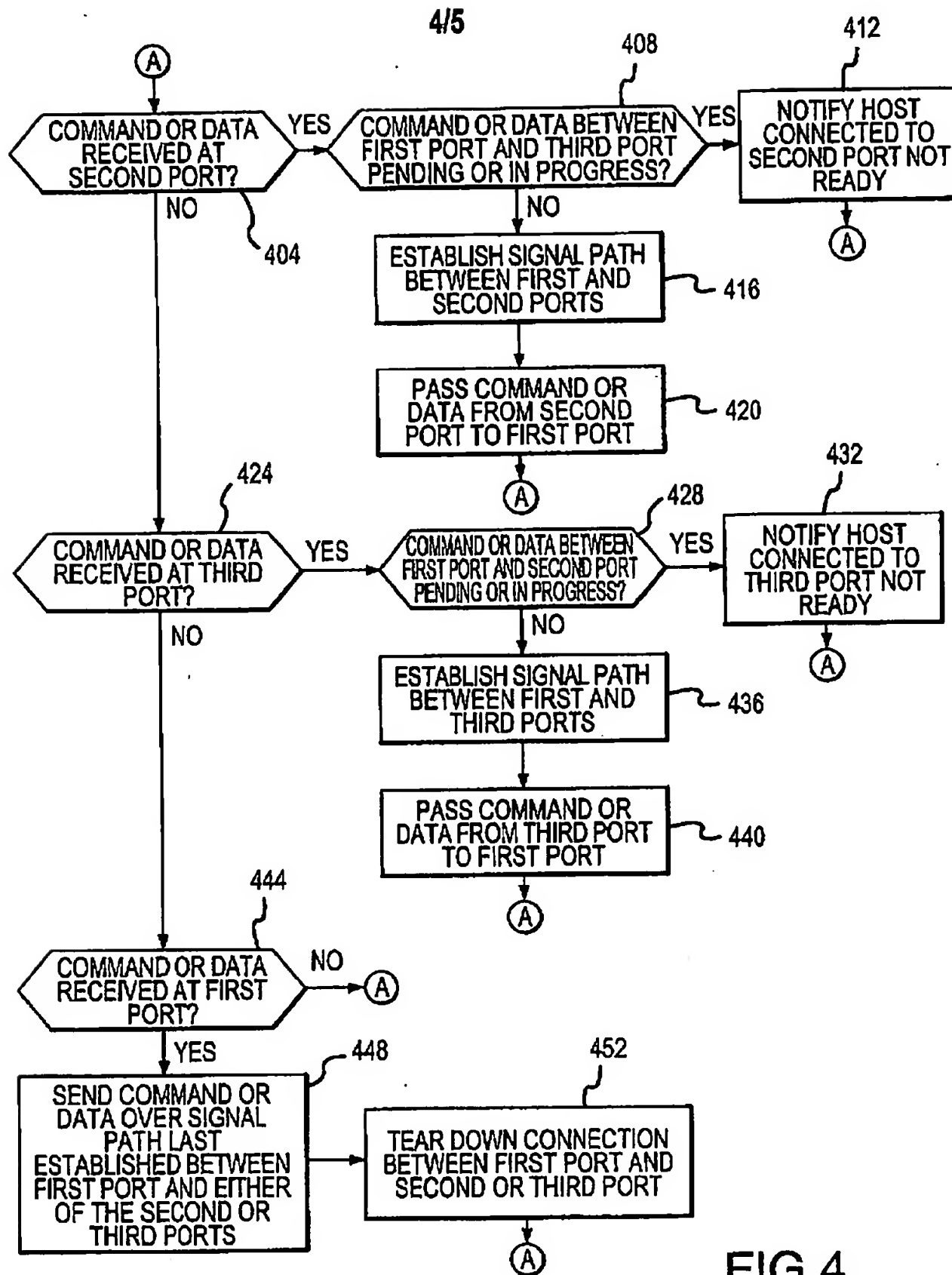


FIG.4

WO 03/091887

PCT/US03/11905

5/5

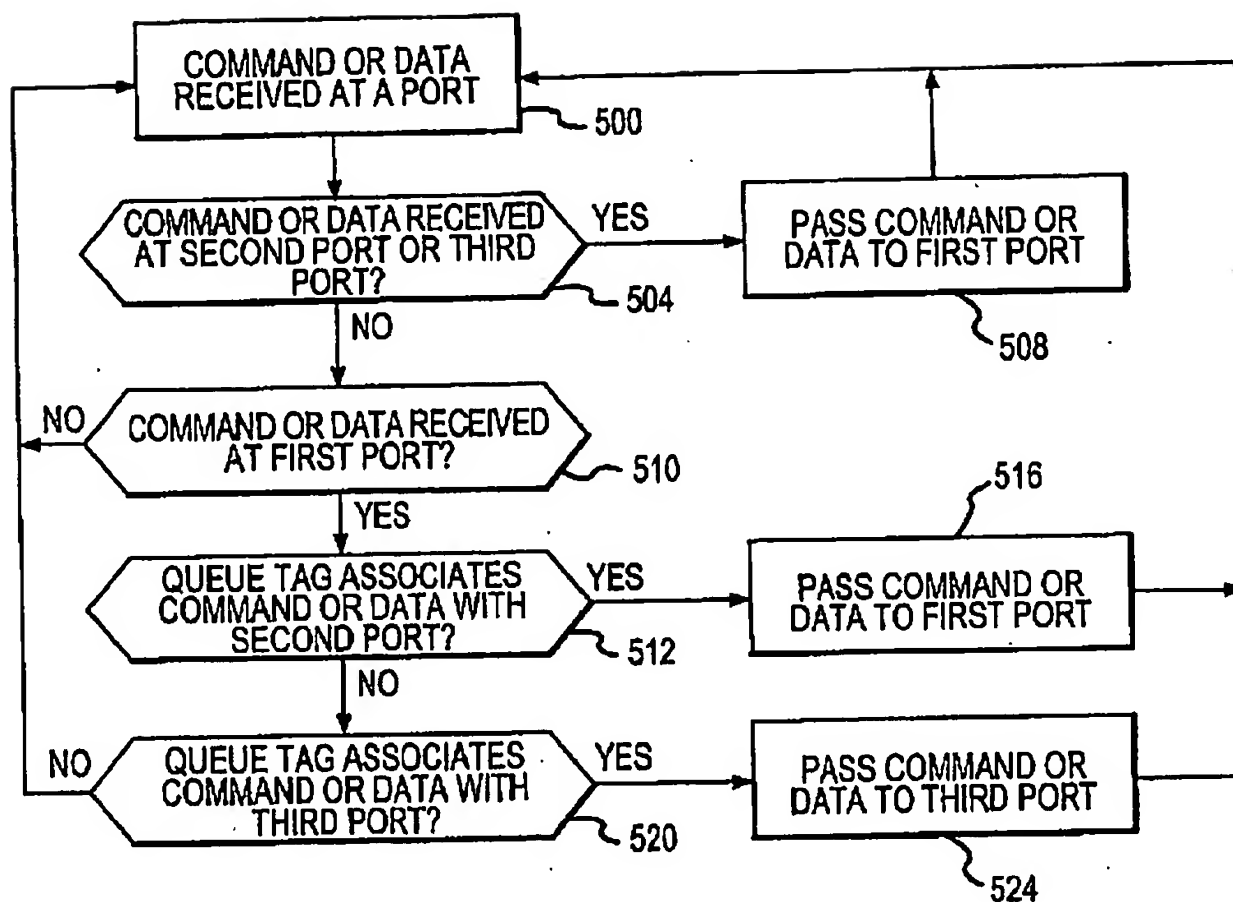


FIG.5

## INTERNATIONAL SEARCH REPORT

International application No.

PCT/US03/11905

## A. CLASSIFICATION OF SUBJECT MATTER

IPC(7) : G06F 13/00

US CL : 710/316

According to International Patent Classification (IPC) or to both national classification and IPC

## B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

U.S. : 710/316, 317, 305, 100; 326/30; 370/351, 422, 423

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)  
Please See Continuation Sheet

## C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	US 5,621,899 A (GAFFORD et al) 15 April 1997 (15.04.1997), columns 9 to 12.	1, 4, 5, 7, 8, 9
X	US 5,245,320 A (BOUTON) 14 September 1993 (14.09.1993), column 4, lines 37 et seq. to col. 7, lines 1-16.	1,2,4,5,7
X,P	US 6,381,642 B1 (O'DONNELL et al) 30 April 2002 (30.04.2002), columns 1-2.	1-10, 15-23
A,P	US 6,449,680 B1 (SANDERS et al) 10 September 2002 (10.09.2002), entire specification.	1-23

☐ Further documents are listed in the continuation of Box C.

See patent family annex.

\* Special categories of cited documents:

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier application or patent published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T"

later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X"

document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when this document is taken alone

"Y"

document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&amp;"

document member of the same patent family

Date of the actual completion of the international search

29 June 2003 (29.06.2003)

Date of mailing of the international search report

22 JUL 2003

Name and mailing address of the ISA/US

Mail Stop PCT, Attn: ISA/US

Commissioner for Patents

P.O. Box 1450

Alexandria, Virginia 22313-1450

Facsimile No. (703)305-3230

Authorized officer

Xuan M. Thai

Telephone No. 703-305-3900

Form PCT/ISA/210 (second sheet) (July 1998)

## INTERNATIONAL SEARCH REPORT

PCT/US03/11905

**Continuation of B. FIELDS SEARCHED Item 3:**

USPAT, USPG-PUB, EPO, JPO, DERWENT, IBMTDB, IEEE

three-way or three-ported switch or connector or interface or adapter, logic circuit, point to point, plural controllers or hosts, RAID, ATA, multiple ports

Form PCT/ISA/210 (second sheet) (July 1998)